REMARKS

Any fees that may be due in connection with the filing of this paper or with this application may be charged to Deposit Account No. 06-1050. If a Petition for Extension of time is needed, this paper is to be considered such Petition.

Claims 74-76, 92-94, 123, 124, 127-133 and 135-139 are pending. Claim 127 is amended herein. Claim 127 is amended to restate the recitation "a selected nucleotide base of the four bases of the nucleic acid" as "a selected nucleotide base of the four nucleic acid nucleotide bases" to more distinctly claim the subject matter. No new matter is added.

REJECTION OF CLAIMS 127-135 AND 137-139 UNDER 35 U.S.C. §112, SECOND PARAGRAPH

Claims 127-135 and 137-139 are rejected under 35 U.S.C. §112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter, because the recitation "the four bases" allegedly lacks antecedent basis in the claim.

This rejection is respectfully traversed.

RELEVANT LAW

35 U.S.C. § 112, second paragraph requires only reasonable precision in delineating the bounds of the claimed invention. Claim language is satisfactory if it reasonably apprises those of skill in the art of the bounds of the claimed invention and is as precise as the subject matter permits. *Shatterproof Glass Corp. v. Libby-Owens Ford Col.*, 758 F.2d 613, 624, 225 USPQ 634, 641 (Fed. Cir.), *cert. dismissed*, 106 S.Ct. 340 (1985). When one skilled in the art would understand all of the language in the claims when read in light of the specification, a claim is not indefinite.

The failure to provide explicit antecedent basis for terms does not always render a claim indefinite. If the scope of a claim would be reasonably ascertainable by those skilled in the art, then the claim is not indefinite. *Ex parte Porter*, 25 USPQ2d 1144, 1145 (Bd. Pat. App. & Inter. 1992). Inherent components of elements recited have antecedent basis in the recitation of the components themselves. For example, the limitation "the outer surface of said sphere" would not require an antecedent recitation that the sphere has an outer surface. See *Bose Corp. v. JBL, Inc.*, 274 F.3d 1354, 1359, 61 USPQ2d 1216, 1218-19 (Fed. Cir 2001).

THE CLAIMS

Claim 127 is directed to an array of nucleic acid probes, where each probe includes a single-stranded portion at one terminus, a double-stranded portion at the opposite terminus,

and a variable nucleotide sequence within the single-stranded portion, where the probes are divided into four subsets and for each subset, a selected nucleotide base of the four nucleic acid nucleotide bases occupies a defined number of positions in each probe and all other nucleotide bases except the selected nucleotide base occupy the remaining positions. Claims 128-135 and 137-139 ultimately depend from claim 127 and are directed to various embodiments thereof.

ANALYSIS

The Examiner alleges that claims 127-135 and 137-139 are indefinite because there is no antecedent basis for the recitation "the four bases" in claim 127.

Applicant respectfully disagrees. Attention is directed to MPEP §2173.05(e) Lack of Antecedent Basis, which recites:

> A claim is indefinite when it contains words or phrases whose meaning is unclear. The lack of clarity could arise where a claim refers to "said lever" or "the lever," where the claim contains no earlier recitation or limitation of a lever and where it would be unclear as to what element the limitation was making reference. Similarly, if two different levers are recited earlier in the claim, the recitation of "said lever" in the same or subsequent claim would be unclear where it is uncertain which of the two levers was intended. A claim which refers to "said aluminum lever," but recites only "a lever" earlier in the claim, is indefinite because it is uncertain as to the lever to which reference is made. Obviously, however, the failure to provide explicit antecedent basis for terms does not always render a claim indefinite. If the scope of a claim would be reasonably ascertainable by those skilled in the art, then the claim is not indefinite. *Ex parte Porter*, 25 USPQ2d 1144, 1145 (Bd. Pat. App. & Inter. 1992) ("controlled stream of fluid" provided reasonable antecedent basis for "the controlled fluid"). Inherent components of elements recited have antecedent basis in the recitation of the components themselves. For example, the limitation "the outer surface of said sphere" would not require an antecedent recitation that the sphere has an outer surface. See *Bose Corp. v. JBL, Inc.*, 274 F.3d 1354, 1359, 61 USPQ2d 1216, 1218-1219 (Fed. Cir. 2001) (holding that recitation of "an ellipse" provided antecedent basis for "an ellipse having a major diameter" because "[t]here can be no dispute that mathematically an inherent characteristic of an ellipse is a major diameter."

In this instance, claim 127 recites in pertinent part:

> for each subset, a selected nucleotide base of the four nucleic acid nucleotide bases occupies a defined number of positions in each probe and all other nucleotide bases except the selected nucleotide base occupy the remaining positions.

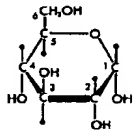
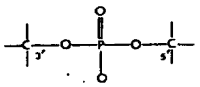
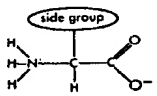
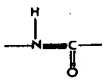
It is respectfully submitted that antecedent basis for "the four bases" is implicit in the term "nucleic acid." It is known to one of skill in the art that a nucleic acid is a polynucleotide, or a polymer consisting of nucleotides (*e.g.*, see Freifelder, *Molecular Biology*, 1983, pages 80-

82). Freifelder teaches, on page 80, that, in addition to a phosphate group attached to the 5' carbon of the sugar by a phosphoester linkage, each nucleotide includes:

> 1. A cyclic five-carbon sugar. This is ribose, in the case of ribonucleic acid (RNA), and deoxyribose, in deoxyribonucleic acid (DNA). The difference in the structure of ribose and 2'-deoxyribose is shown in Figure 3.5. Note that they differ only in the absence of a 2'-OH group in deoxyribose, a difference that makes DNA chemically more stable than RNA, as will be seen later.
>
> 2. A purine or pyrimidine base attached to the 1'-carbon atom of the sugar by an *N*-glycosidic bond. The bases, which are shown in Figure 3-6, are the purines, adenine (A) and guanine (G), and the pyrimidines, cytosine (C), thymine (T), and uracil (U). DNA and RNA contain A, G and C; however, T is found only in DNA and U is found only in RNA.

Watson (*Molecular Biology of the Gene*, 2nd ed, 1970, Table 3-4, page 86) teaches that a DNA molecule and an RNA molecule each include four nucleotide bases as monomers of the polynucleotide:

TABLE 3–4 *Structural organization of several important biological macromolecules*

Macromolecule	Monomeric units	Number of different monomers	General monomer formula	Fixed or irregular chain length	Linkage between monomers
Glycogen (a polysaccharide)	Glucose	One		Indefinite—may be > 1000	1–4-Glycosidic linkage
DNA (deoxyribonucleic acid)	Deoxynucleotides	Four: deoxyadenylate deoxyguanylate deoxythymidylate deoxycytidylate	Purine-deoxyribose-P (or pyrimidine-deoxyribose-P)	Genetically fixed—may be > 10^7	3'–5'-Phosphodiester linkage 
RNA (ribonucleic acid)	Ribonucleotides	Four: adenylate guanylate uridylate cytidylate	Purine-ribose-P (or pyrimidine-ribose-P)	Genetically fixed, often > 3000	3'–5'-Phosphodiester linkage
Protein	L-Amino acids	Twenty: glycine, alanine, serine, etc.		Genetically fixed, usually varies between 100 and 1000	Peptide linkage 

Zubay teaches that both DNA and RNA contain four different nucleotides, each of which contains a nitrogenous base (Zubay, *Biochemistry*, (1983), page 661). Zubay teaches that the four bases found in DNA are adenine, thymine, guanine and cytosine, and that the four bases in RNA are adenine, uracil, guanine and cytosine (Zubay, *Biochemistry*, (1983), pages 662-663). Zubay teaches that the four deoxyribonucleotides in a DNA molecule are deoxyadenosine-5'-phosphate, deoxythymidine-5'-phosphate, deoxyguanosine-5'-phosphate and deoxycytidine-5'-phosphate and the four ribonucleotides in an RNA molecule are

adenosine-5'-phosphate, uridine-5'-phosphate, cytidine-5'-phosphate and guanine-5'-phosphate (*Id.*).

Thus, one of skill in the art would recognize that nucleic acid molecules include as inherent components the bases that constitute the polynucleotide structure. Specifically, one of skill in the art would understand a DNA molecule to include the four nucleotide bases adenine, cytosine, guanine and thymine and an RNA molecule to include the four nucleotide bases cytosine, uracil, guanine and adenine. As discussed above, inherent components of elements recited in the claims have antecedent basis in the recitation of the components themselves. Thus, the recitation "the four bases" does not require an antecedent recitation that a nucleic acid contains four nucleotide bases, since the four nucleotide bases are inherent components of the nucleic acid.

In addition, Applicant respectfully submits that one of skill in the art, in light of the teachings of the specification and what is known in the art, would understand the meaning of the recitation "a selected nucleotide base of the four bases of the nucleic acid" to refer to one of deoxyadenosine-5'-phosphate, deoxythymidine-5'-phosphate, deoxyguanosine-5'-phosphate or deoxycytidine-5'-phosphate for a DNA molecule and one of adenosine-5'-phosphate, uridine-5'-phosphate, cytidine-5'-phosphate and guanine-5'-phosphate for an RNA molecule. Thus, the scope of the claim would be reasonably ascertainable by those skilled in the art. Therefore, the recitation "a selected nucleotide base of the four bases of the nucleic acid" does not render the claim indefinite.

THE REJECTION OF CLAIMS 127-133 AND 135-139 UNDER 35 U.S.C. §102(b)

Claims 127-133 and 135-139 are rejected under 35 U.S.C. §102(b) as anticipated by Holmes *et al.* (WO 90/06045, published 14 June 1990) because Holmes *et al.* allegedly discloses every element of the claimed array. The Examiner alleges that the non-hybridized target of Holmes *et al.* is a "variable region," the oligo-dT is a "double-stranded portion," that a biotinylated nucleotide at the end or an incorporated ddNTP is a "selected nucleic acid base that occupies a defined number of positions" and "non-biotinylated A, T, C and G" are "all other bases" that occupy the remaining positions, and that page 15 discloses dividing the arrays into for subsets. This rejection is respectfully traversed.

RELEVANT LAW

Anticipation requires the disclosure in a single prior art reference of each element of the claim under consideration. *In re Spada*, 15 USPQ2d 1655 (Fed. Cir, 1990), *In re Bond*, 15 USPQ 1566 (Fed. Cir. 1990), *Soundscriber Corp. v. U.S.*, 360 F.2d 954, 148 USPQ 298, 301, adopted 149 USPQ 640 (Ct. Cl.) 1966. See, also, *Richardson v. Suzuki Motor Co.*, 868 F.2d

1226, 1236, 9 USPQ2d 1913,1920 (Fed. Cir.), *cert. denied*, 110 S.Ct. 154 (1989). "[A]ll limitations in the claims must be found in the reference, since the claims measure the invention". *In re Lang*, 644 F.2d 856, 862, 209 USPQ 288, 293 (CCPA 1981). Moreover it is incumbent on Examiner to identify wherein each and every facet of the claimed invention is disclosed in the reference. *Lindemann Maschinen-fabrik Gmbh v. American Hoist and Derrick Co.*, 730 F.2d 1452, 221 USPQ 481 (Fed. Cir. 1984). Further, the reference must describe the invention as claimed sufficiently to have placed a person of ordinary skill in the art in possession of the invention. An inherent property has to flow naturally from what is taught in a reference *In re Oelrich*, 666 F.2d 578, 581, 212 USPQ 323, 326 (CCPA 1981).

THE CLAIMS

See related section above.

Disclosure of Holmes *et al.*

Holmes *et al.* discloses magnetic particles that have a plurality of single-stranded oligonucleotide probes that are either directly attached to the magnetic particle or attached via a double-stranded piece of DNA (page 2, lines 9-22). The probes include an oligo-dT sequence that will hybridize with poly A tails "universally present on native eukaryotic mRNA" and can include specific DNA sequences that hybridize with specific sequences in target RNA and single-stranded DNA molecules (page 2, lines 9-22). In one embodiment, a DNA molecule having a sequence complementary to a known sequence of a target nucleic acid molecule is hybridized to the probe (having an oligo-dT sequence) via a poly-dA tail on the DNA sequence, and a different labeled "probe" is hybridized to a different sequence of the nucleic acid molecule, forming a ternary complex (page 13, lines 1-11). Holmes *et al.* discloses standard Sanger sequencing reactions (page 14, lines 2-36). Holmes *et al.* discloses a method of sequencing single-stranded nucleic acids that includes dividing its magnetic particles that include single-stranded DNA or RNA oligonucleotide to be sequenced into four aliquots and adding to each aliquot a polymerase, mixed nucleotide triphosphates and a different dideoxynucleoside triphosphate for each aliquot (page 15, lines 15-37).

Differences between the claimed subject matter and the disclosure of Holmes *et al.*

Claim 127 is directed to an array of nucleic acid probes, where each probe includes a single-stranded portion at one terminus, a double-stranded portion at the opposite terminus, and a variable nucleotide sequence within the single-stranded portion. The probes of the array of claim 127 are divided into four subsets and for each subset, a selected nucleotide base of the four nucleic acid nucleotide bases occupies a defined number of positions in each

probe and all other nucleotide bases except the selected nucleotide base occupy the remaining positions.

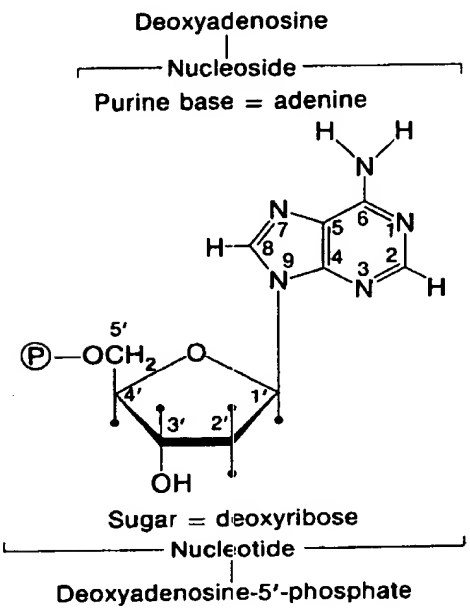
The Examiner alleges that the oligonucleotide to be sequenced in Holmes *et al.* is "a single-stranded region having a variable region" and that the oligo-dT portion in Holmes *et al.* is "a double-stranded portion." The Examiner alleges that Holmes *et al.* discloses dividing an array of nucleic acid into four subsets for sequencing, where for each subset, a dideoxy analogue (ddNTP, such as ddATP) is attached by the DNA polymerase. The Examiner alleges that this ddNTP can be viewed as the element "a selected nucleic acid base" that "occupies a defined number of positions" [one] and that the other bases of the oligonucleotide to be sequenced in Holmes *et al.* can be viewed as the element "all other bases except the selected nucleotide base occupy the remaining positions." Applicant respectfully disagrees.

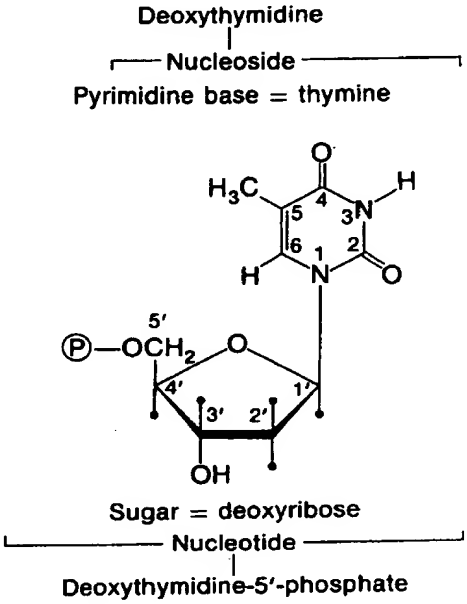
1. ddNTP is not one of the four nucleotide bases of the nucleic acid

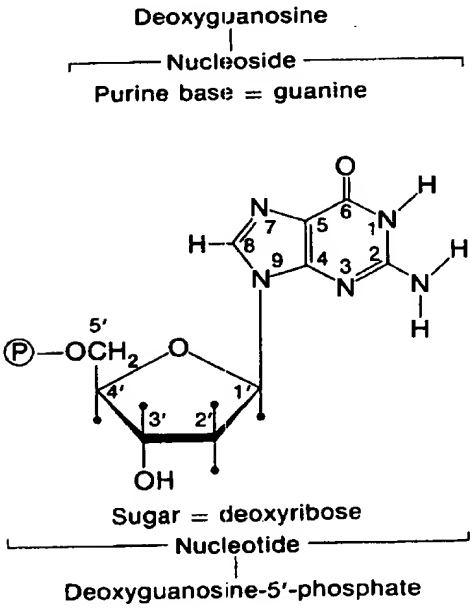
At the section cited by the Examiner (page 14, lines 2-20), Holmes *et al.* discloses Sanger sequencing, which incorporates a terminal dideoxynucleoside triphosphate in the probe to produce a series of labeled DNA strands having different chain lengths and ending with a particular dideoxynucleotide triphosphate base to produce a set of nested fragments. Holmes *et al.* discloses incorporating a dideoxynucleotide triphosphate analogue to block further growth of the new chain because the dideoxynucleotide triphosphate analogue lacks the 3'-hydroxyl terminus needed to form the next phosphodiester bond (page 14, lines 10-13). This nested set of fragments with the dideoxynucleotide triphosphate analogue at the 3' end is used for sequencing.

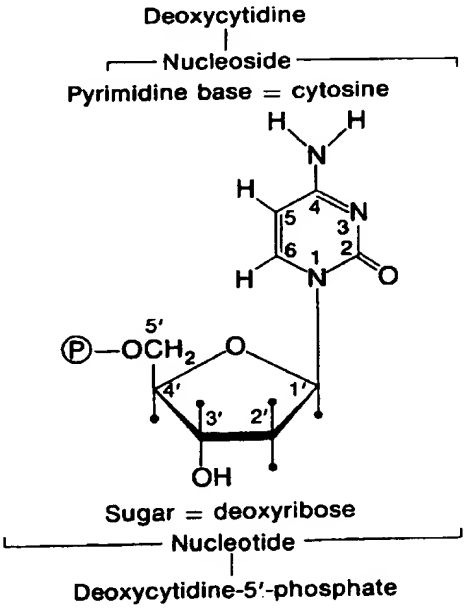
The instant claims recite that a nucleotide base is selected from one of the four nucleic acid nucleotide bases and occupies a defined number of positions in each probe. As discussed above, it is recognized by one of skill in the art that that a nucleic acid molecule includes four nucleotide bases, and that each nucleotide of a nucleic acid includes a phosphate group attached to the 5' carbon of the sugar by a phosphoester linkage, a cyclic five-carbon sugar (ribose in RNA and deoxyribose in DNA), and a purine or pyrimidine base attached to the 1'-carbon atom of the sugar by an N-glycosidic bond (*e.g.*, see Freifelder, *Molecular Biology*, 1983, page 80). The Examiner alleges that the ddNTP of Holmes *et al.* is a nucleotide base because the absence of the 3'-hydroxyl does not alter the base. Applicant respectfully disagrees. Freifelder teaches that a *nucleoside* is a purine or pyrimidine base linked to a sugar (ribose or deoxyribose) and that a nucleotide includes the phosphate moiety and thus a nucleotide is a nucleoside phosphate (*Id.*) Zubay teaches that the ribonucleotides of an RNA molecule are adenosine-5'-phosphate, uridine-

5'-phosphate, cytidine-5'-phosphate and guanine-5'-phosphate and that the deoxynucleotides of a DNA molecule are deoxyadenosine-5'-phosphate, deoxythymidine-5'-phosphate, deoxyguanosine-5'-phosphate and deoxycytidine-5'-phosphate (Zubay, *Biochemistry*, pages 661-663 (1983). This is clearly shown in Figure 18-1 of Zubay (see page 662), which is reproduced below:









None of adenosine-5'-phosphate, uridine-5'-phosphate, cytidine-5'-phosphate, guanine-5'-phosphate, deoxynucleotides deoxyadenosine-5'-phosphate, deoxythymidine-5'-phosphate,

deoxyguanosine-5'-phosphate or deoxycytidine-5'-phosphate is a dideoxynucleotide triphosphate. One of skill in the art, as evidenced by the teachings of Holmes *et al.*, recognizes that a dideoxynucleotide triphosphate (ddNTP) is an analogue of DNA nucleotide in which the 3'-OH group is replaced with a hydrogen, which prevents DNA polymerase from adding another nucleotide to a growing DNA strand (see page 14, lines 5-19).

Holmes *et al.* discloses on page 15 that its magnetic particles include a single-stranded oligonucleotide to be sequenced and that these particles are divided into four aliquots. To each of these aliquots a different dideoxynucleoside triphosphate is added to the terminus to created a nested set of differently terminated fragments for sequencing. Holmes *et al.* does not disclose an array of probes divided into four subarrays, where for each subarray, one of the four nucleic acid nucleotide bases is selected to occupy a defined number of positions and the other positions are occupied by nucleotide bases other than the selected nucleotide base. Holmes *et al.* does not disclose an array of DNA probes where the array is divided into four subsets and for each subset, one of deoxyadenosine-5'-phosphate, deoxythymidine-5'-phosphate, deoxyguanosine-5'-phosphate or deoxycytidine-5'-phosphate is selected and occupies a defined number of positions in each probe and all other deoxynucleotide bases except the selected base occupy the remaining positions. Holmes *et al.* does not disclose an array of RNA probes where the array is divided into four subsets and for each subset, one of adenosine-5'-phosphate, uridine-5'-phosphate, cytidine-5'-phosphate or guanine-5'-phosphate is selected and occupies a defined number of positions in each probe and all other nucleotide bases except the selected nucleotide base occupy the remaining positions. Hence, Holmes *et al.* does not disclose every element of claim 127. Therefore, because Holmes *et al.* does not disclose every element of claim 127, Holmes *et al.* does not anticipate claims 127-133 and 135-139.

REBUTTAL TO THE EXAMINER'S ARGUMENTS

1. No Evidence Provided in Previous Response

The Examiner alleges that Applicant did not provide any evidence that one of skill in the art would recognize that the four nucleotide bases of a nucleic acid include RNA nucleotides (A, U, C, G) or DNA nucleotides (A, T, C, G), and thus the Examiner deemed the assertion arguments of counsel. Applicant respectfully disagrees. Applicant directed the Examiner to the teachings of Zubay (*Biochemistry*, (1983), pages 661-663), a copy of which was submitted with the response filed November 16, 2005. This reference was made of record,

as evidenced by its entry into PAIR (see the entry dated November 16, 2005 entitled "Applicant Arguments/Remarks Made in an Amendment," pages 13-17). As discussed above, Zubay teaches that DNA and RNA are polynucleotides and that each contains four different nucleotides, that the four nucleotide bases found in DNA are deoxyadenosine-5'-phosphate, deoxythymidine-5'-phosphate, deoxyguanosine-5'-phosphate and deoxycytidine-5'-phosphate, and that the four nucleotide bases in RNA are adenosine-5'-phosphate, uridine-5'-phosphate, cytidine-5'-phosphate and guanine-5'-phosphate. Thus, the teaching of Zubay is evidence that was placed in the record and provided to the Examiner to support the argument that one of skill in the art would recognize that the four nucleotide bases of a DNA molecule are deoxy-adenosine-5'-phosphate, deoxythymidine-5'-phosphate, deoxyguanosine-5'-phosphate and deoxycytidine-5'-phosphate and the four nucleotide bases for an RNA molecule are adenosine-5'-phosphate, uridine-5'-phosphate, cytidine-5'-phosphate and guanine-5'-phosphate.

Applicant provides herein additional evidence in support of this argument, as discussed in detail above under the rejection under 35 U.S.C. §112, second paragraph. For example, Watson teaches that a DNA molecule includes the deoxynucleotides deoxyadenylate, deoxyguanylate, deoxythymidylate and deoxycytidylate, and that an RNA molecule includes the ribonucleotides adenylate, guanylate, uridylate and cytidylate (Watson, *Molecular Biology of the Gene* (2nd ed., 1970), page 86). Freifelder teaches that a DNA molecule includes nucleotides that include the bases adenine, guanine, cytosine and thymine, where an RNA molecule includes nucleotides that include the bases adenine, guanine, cytosine and uracil (Freifelder, *Molecular Biology* (1983), pages 80-82). Garrett & Grisham teaches that an RNA molecule includes nucleotides that include the bases adenine, cytosine, guanine and uracil and that a DNA molecule includes nucleotides that include the bases thymine, guanine, cytosine and adenine (Garrett & Grisham, *Biochemistry* (1995), page 189). When referring to the bases that constitute a nucleic acid molecule, one of skill in the art generally refers to the bases of the polynucleotide. For example, Garrett & Grisham teaches on page 190, lines 1-16 that:

> The only significant variation that commonly occurs in the chemical structure of nucleic acids is the nature of the base at each nucleotide position. These bases are not part of the sugar-phosphate backbone but instead serve as distinctive side chains, like the R groups attached to a polypeptide backbone. They give the polymer its unique identity. A simple notation of these structures is merely to list the order of the bases in the polynucleotide using single capital letters – A, G, C and U (or T). Occasionally, a lowercase "p" is written between each successive base to indicate the phosphodiester bridge, as in GpApCpGpUpA. A "p" preceding the sequence indicates that the nucleic acid carries a PO_4 on its

5'-end, as in pGpApCpGpUpAp; a "p" terminating the sequence connotes the presence of a phosphate on the 3'-OH end, as in GpApCpGpUpAp.

A more common method of representing nucleotide sequences is to omit the "p" and write only the order of the bases, such as GACGUA. This notation assumes the presence of the phosphodiesters joining adjacent nucleotides.

Bohinski teaches that RNA and DNA are composed of monomeric units called nucleotides and that bases found in a DNA molecule are adenine, guanine, cytosine and thymine, while the bases found in an RNA molecule are adenine, guanine, uracil and cytosine (Bohinski, *Modern Concepts in Biochemistry*, 4th ed., (1983), pages 164-165). Thus, one of skill in the art would recognize that the four nucleotide bases of a nucleic acid include RNA nucleotides (A, U, C, G) or DNA nucleotides (A, T, C, G)

2. Five Bases Render The Argument Not Commensurate in Scope of the Claims

The Examiner states that Applicant lists five nucleotide bases, A, U, T, C and G. The Examiner alleges that these five nucleotide bases meet the limitations of the claims and alleges that because there are five nucleotide bases, the arguments are not commensurate in scope with the claimed four nucleotide bases. Applicant respectfully disagrees. Claim 127 is directed to an array of nucleic acid probes, where the array includes as an element that the probes are divided into four subsets and for each subset, a selected nucleotide base of the four nucleic acid bases occupies a defined number of positions in each probe and all other nucleotide bases except the selected nucleotide base occupy the remaining positions. As discussed above, one of skill in the art recognizes the term "nucleic acid" to include deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) (*e.g.*, see Freifelder, *Molecular Biology*, 1983, pages 80-82). Further, as discussed above, one of skill in the art knows that both DNA and RNA contain four different nucleotides, that the four nucleotide bases found in DNA are deoxyadenosine-5'-phosphate, deoxythymidine-5'-phosphate, deoxyguanosine-5'-phosphate and deoxycytidine-5'-phosphate and the four nucleotide bases for an RNA molecule are adenosine-5'-phosphate, uridine-5'-phosphate, cytidine-5'-phosphate and guanine-5'-phosphate (Zubay, Biochemistry, (1983), pages 661-663). Thus, when the nucleic acid in claim 127 is DNA, the four nucleotide bases are deoxyadenosine-5'-phosphate, deoxythymidine-5'-phosphate, deoxyguanosine-5'-phosphate and deoxycytidine-5'-phosphate, and when the nucleic acid in claim 127 is RNA, the four nucleotide bases are adenosine-5'-phosphate, uridine-5'-phosphate, cytidine-5'-phosphate and guanine-5'-phosphate. Hence, Applicant respectfully submits that the arguments are commensurate in scope with the claimed four nucleic acid nucleotide bases.

* * *

In view of the above, examination of the application on the merits and allowance is respectfully requested.

Respectfully submitted,

Stephanie Seidman
Reg. No. 33,779

Attorney Docket No. 17120-002007 / 2401G
Address all correspondence to:
Stephanie Seidman
Fish & Richardson P.C.
12390 El Camino Real
San Diego, California 92130
Telephone: (858) 678-5070
Facsimile: (202) 626-7796
email: seidman@fr.com

RESPONSE UNDER 37 CFR §1.116
--EXPEDITED PROCEDURE--
EXAMINING GROUP 1600

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant : Cantor *et al.*
Serial No. : 09/030,571
Conf. No. : 7542
Filed : February 24, 1998
Title : **POSITIONAL SEQUENCING BY HYBRIDIZATION**

Art Unit : 1634
Examiner : B. Forman
Customer No.: 20985

MAIL STOP AF
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

SUPPORTING DOCUMENTS

1. Freifelder, *Molecular Biology*, 1983, pages 80-82.
2. Zubay, *Biochemistry*, 1983, pages 661-663.
3. Watson, *Molecular Biology of the Gene*, 2nd ed, 1970, page 86.
4. Garrett & Grisham, *Biochemistry* (1995), page 189-190.
5. Bohinski, *Modern Concepts in Biochemistry*, 4th ed., (1983), pages 164-165.

Molecular Biology

A Comprehensive Introduction to Prokaryotes and Eukaryotes

DAVID FREIFELDER
University of California, San Diego
University of Alabama
Formerly of Brandeis University

Jones and Bartlett Publishers, Inc.
BOSTON PORTOLA VALLEY

Editorial offices: Jones and Bartlett Publishers, Inc., 30 Granada Court, Portola Valley, CA 94025.

Sales and customer service offices: Jones and Bartlett Publishers, Inc., 20 Park Plaza, Boston, MA 02116.

Library of Congress Cataloging in Publication Data

Freifelder, David Michael, 1935–
Molecular biology.

Includes bibliographies and index.
1. Molecular biology I. Title.
QH506.F73 1983 574.8'8 82–17006
ISBN 0–86720–012–X

ISBN 0-86720-012-X

Publisher Arthur C. Bartlett

Book and Cover Design Hal Lockwood

Illustrator Donna Salmon, Assisted by Cyndie Clark-Huegel, Evanell Towne, John and Judy Waller, Dorothy Beebe, Kelly Solis-Navarro, and Brenda Booth

Manuscript Editor Kirk Sargent

Production Bookman Productions

Composition Typothetae

Printer and Binder Halliday Litho

Printed in the United States of America
Printing Number (last digit) 10 9 8 7 6 5 4 3 2

same or different polypeptide chain. This is important in determining the three-dimensional structure of a protein, as will be examined in Chapter 5.

Many proteins contain metal ions engaged in coordination complexes with groups in the side chains. Common metal ions are Zn^{2+}, Cu^{2+}, and Fe^{3+}, which are often bound to histidine and glutamic acid.

Nucleic Acids

A nucleic acid is a polynucleotide—that is, a polymer consisting of nucleotides. Each nucleotide has the three following components (Figure 3-5):

1. A cyclic five-carbon sugar. This is ribose, in the case of ribonucleic acid (RNA), and deoxyribose, in deoxyribonucleic acid (DNA). The difference in the structure of ribose and 2′-deoxyribose is shown in Figure 3-5. Note that they differ only in the absence of a 2′-OH group in deoxyribose, a difference that makes DNA chemically more stable than RNA, as will be seen later.
2. A purine or pyrimidine base attached to the 1′-carbon atom of the sugar by an *N*-glycosylic bond. The bases, which are shown in Figure 3-6, are the purines, adenine (A) and guanine (G), and the pyrimidines, cytosine (C), thymine (T), and uracil (U). DNA and RNA contain A, G, and C; however, T is found only in DNA and U is found only in RNA. There are rare exceptions to this rule—namely, T is present in some tRNA molecules and there are a few phages whose DNA exclusively contains U rather than T.
3. A phosphate attached to the 5′ carbon of the sugar by a phosphoester linkage. This phosphate is responsible for the strong negative charge of both nucleotides and nucleic acids.

A base linked to a sugar is called a **nucleoside**; thus *a nucleotide is a nucleoside phosphate*. The terminology used to describe nucleic acid components is listed in Table 3-1.

The nucleotides in nucleic acids are covalently linked by a second phosphoester bond that joins the 5′-phosphate of one nucleotide and the 3′-OH group of the adjacent nucleotide (Figure 3-7). Thus the phosphate is esterified to both the 3′- and 5′-carbon atoms; this unit is often called a **phosphodiester group.**

The purine and pyrimidine bases are not engaged in any covalent bonds to each other. Thus, a polynucleotide consists of an alternating sugar-phosphate backbone having one 3′-OH terminus and one 5′-phosphate (5′-P) terminus. In the laboratory, polynucleotides can be prepared that have 3′-P and 5′-OH termini, but such molecules do not occur naturally.

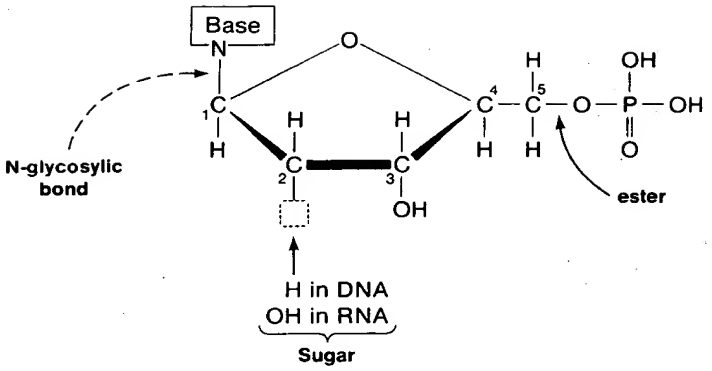
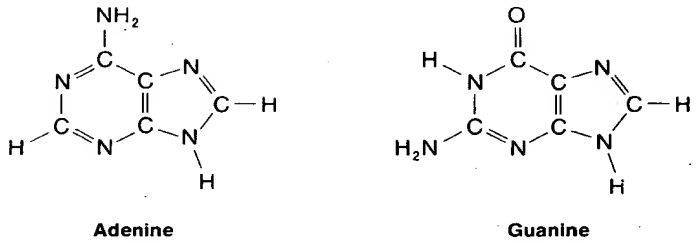


Figure 3-5
Structure of a mononucleotide. The carbon atoms in the sugar ring are numbered.



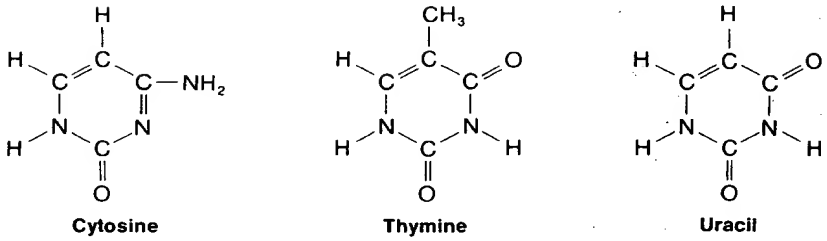


Figure 3-6
The bases found in nucleic acids. The weakly charged groups are shown in red.

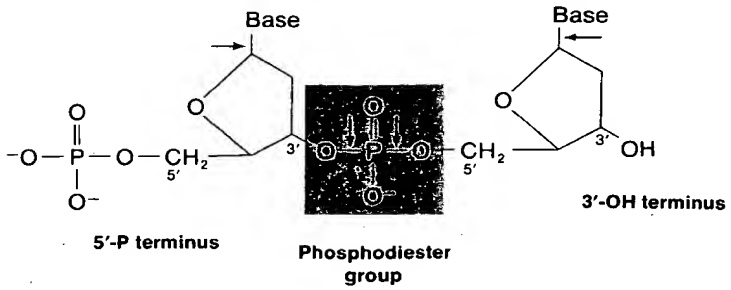


Figure 3-7
The structure of a dinucleotide. The vertical arrows show the bonds in the phosphodiester group about which there is free rotation. The horizontal arrows indicate the *N*-glycosylic bond about which the base can freely rotate. A polynucleotide would consist of many nucleotides linked together by phosphodiester bonds.

A typical RNA molecule is the single-stranded polyribonucleotide that has just been described. However, except in unusual cases, DNA contains two polydeoxynucleotide strands wrapped around one another to form a double-stranded helix. This remarkable structure will be described in the following chapter.

Table 3-1
Nucleic acid nomenclature

Base	*Nucleoside*[1]	*Nucleotide*[2]
	Purines (Pu)	
Adenine (A)	Adenosine (rA)	Adenylic acid, or adenosine monophosphate (AMP)
	Deoxyadenosine (dA)	Deoxyadenylic acid, or deoxyadenosine monophosphate (dAMP)
Guanine (G)	Guanosine[3] (rG)	Guanylic acid, or guanosine monophosphate (GMP)
	Deoxyguanosine (dG)	Deoxyguanylic acid, or deoxyguanosine monophosphate (dGMP)
	Pyrimidines (Py)	
Cytosine (C)	Cytidine (rC)	Cytidylic acid, or cytidine monophosphate (CMP)
	Deoxycytidine (dC)	Deoxycytidylic acid, or deoxycytidine monophosphate (dCMP)
Thymine (T)	Thymidine[4] (dT)	Thymidylic acid, or thymidine monophosphate (TMP)
Uracil (U)	Uridine[5] (rU)	Uridylic acid, or uridine monophosphate (UMP)

[1]Note that the names of purine nucleosides end in -osine and the names of pyrimidine nucleosides end in -idine.

[2]Note that each nucleotide has two names for the same substance.

[3]Guanosine should not be confused with guanidine, which is not a nucleic acid base.

[4]Thymidine is the deoxy- form. The ribo- form, ribosylthymine, is not generally found in nucleic acids.

[5]Uridine is the ribo- form. Deoxyuridine is not commonly found, although deoxyuridylic acid is on the pathway for synthesis of thymidylic acid—i.e., deoxyuridylic acid is methylated to yield thymidylic acid.

BIOCHEMISTRY

Coordinating Author GEOFFREY ZUBAY
COLUMBIA UNIVERSITY

ADDISON-WESLEY PUBLISHING COMPANY

READING, MASSACHUSETTS ▲ MENLO PARK, CALIFORNIA ▲ LONDON ▲ AMSTERDAM ▲ DON MILLS, ONTARIO ▲ SYDNEY

Sponsoring Editor	*Bob Rogers*
Production Editor	*Marcia Mirski*
Copy Editor	*James K. Madru*
Text Designer	*Vanessa Piñeiro*
Illustrator	*Illustration Concepts, Michael Ockler*
Cover Designer and Illustrator	*Hannus Design Associates, Richard Hannus*
Art Coordinator	*Kristin Belanger*
Production Manager	*Karen M. Guardino*
Production Coordinator	*Peter Petraitis*

The text of this book was composed in Trump by York Graphic Services.

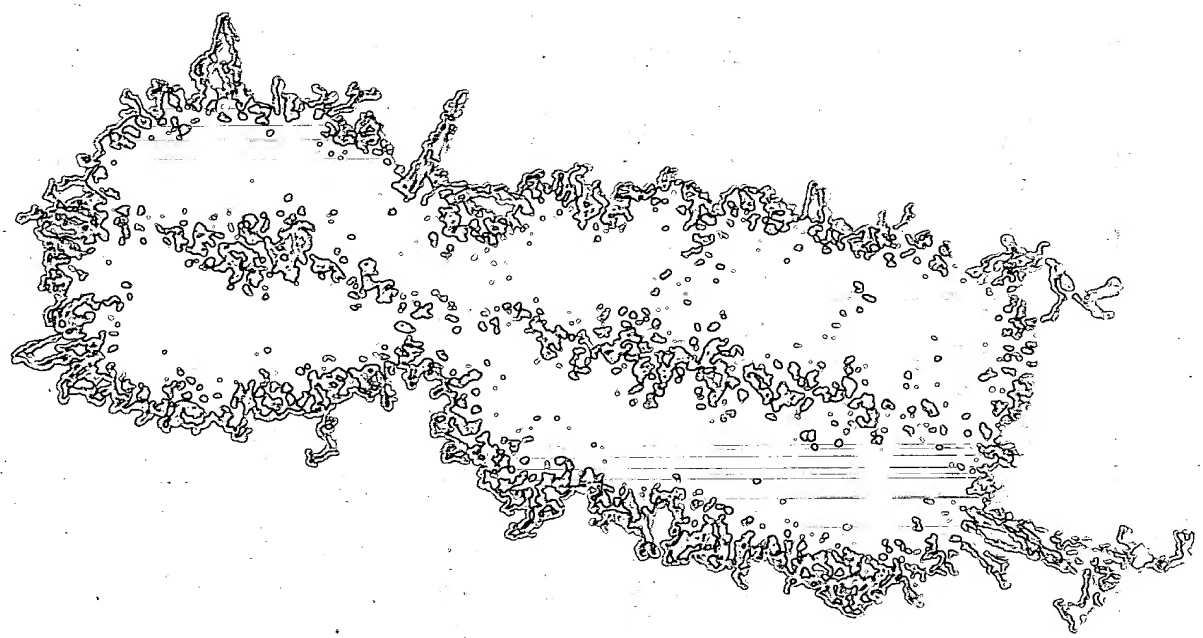
Illustrations rendered and copyrighted by Irving Geis: Figures 1.1, 1.2, 1.3, 1.6, 1.7, 1.8, 1.9, 1.10, 1.11, 1.13, 1.14, 1.15, 1.16, 1.17, 1.18, 1.19, 1.20, 1.21, 1.23, 1.26, 1.27, 1.28, 1.29, 1.32, 1.33, 3.4, 3.7, 3.16, 3.17(a), 3.43, 3.44, 3.45, 3.46, 3.47, 3.54(a), 3.58(lower half), 3.59, 4.7, 4.8, 4.9, 4.10, 4.15, 4.21, 10.9, 12 opener, 18.16, 18.35(b).

Library of Congress Cataloging in Publication Data

Zubay, Geoffrey L.
Biochemistry.

Includes bibliographies and index.
1. Biological chemistry. I. Title.
QP514.2.Z83 1983 574.19'2 82-18502
ISBN 0-201-09091-0

ISBN 0-201-09091-0
ABCDEFGHIJ-DO-89876543



18

STRUCTURE OF NUCLEIC ACIDS AND NUCLEOPROTEINS

Nucleic acids are long-chain polymers composed of nucleotides. The sequence of nucleotides is the repository of all genetic information carried by chromosomes. Despite this, not all nucleic acid is informational, nor is all the informational nucleic acid found in the chromosome. Examples of non-informational nucleic acid include ribosomal RNA and centromeric DNA, whose functions are primarily structural. Examples of informational nucleic acids not found in chromosomes include nucleic acids of mitochondria, chloroplasts, plasmids, and viruses. Most of the chapters in Part IV and Chapters 27 and 28 in Part V are devoted to explaining the ways in which nucleic acids are replicated and transmit their genetic information for use in the cell. In this chapter the focus is on the basic structural properties of nucleic acids in the free-solution state and as they exist in protein complexes in cells.

Electron micrograph of a human chromosome in late-prophase. (Magnification 21,000×.) The chromosome consists of two identical chromatids united at their centromeres. The chromatin consists primarily of a complex of DNA and histone (see text). It is still a mystery as to what forces cause the condensation of the nucleohistone into this highly condensed form. (Micrograph obtained from Gunter F. Bahr, M.D.)

NUCLEOTIDES, THE BUILDING BLOCKS OF NUCLEIC ACIDS

There are two chemically different types of nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). Both DNA and RNA contain four different nucleotides. Each *nucleotide* contains a nitrogenous base known as a *purine* or a *pyrimidine*; a sugar, *ribose* in RNA; *deoxyribose* in DNA; and a phosphoryl group. The nucleotide may be converted to a *nucleoside* by removal of the phosphate. The primary structure of the four commonly occurring deoxyribonucleotides found in DNA are shown in Figure 18-1.

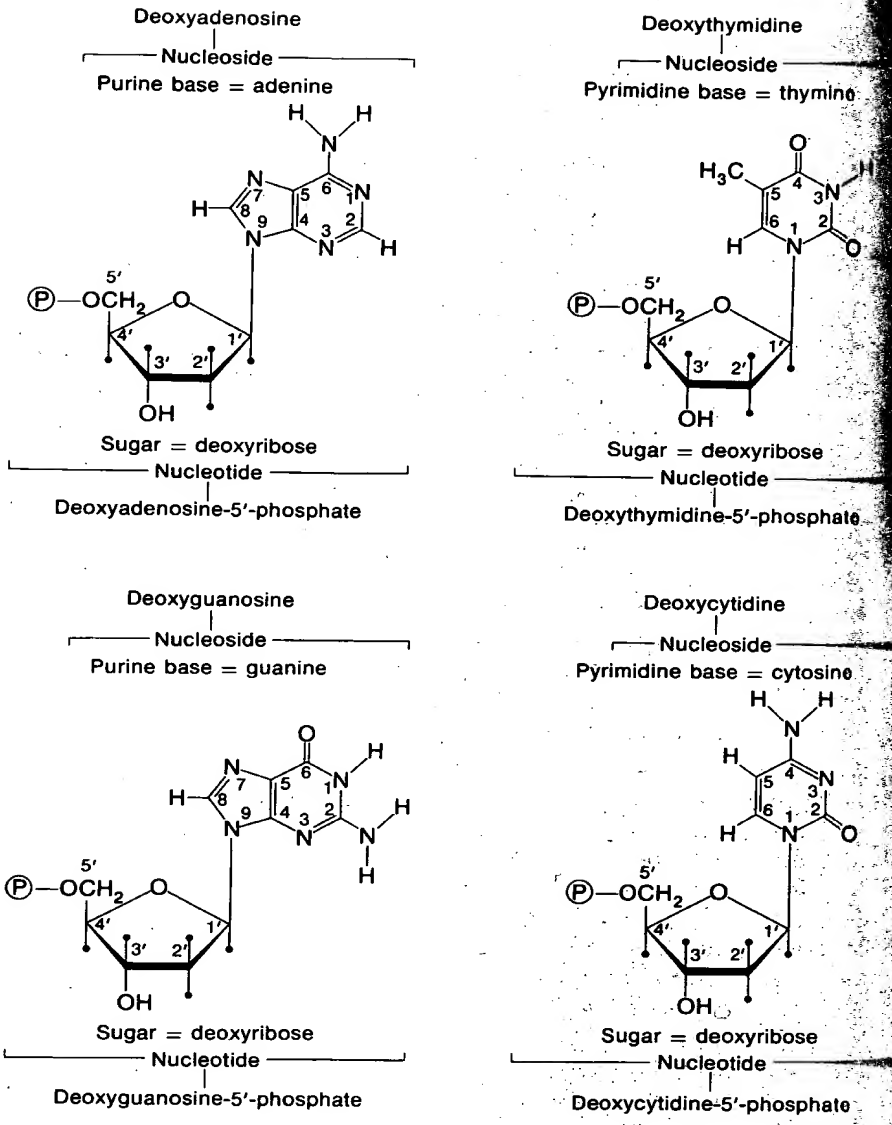


Figure 18-1
Structure of four deoxyribonucleotides found in DNA. Note the numbering system for the carbon and nitrogen atoms of the purine and pyrimidine bases. The carbon atoms of the sugars are usually given prime designations. Note that the nitrogen bases are cis relative to the C-5′ and trans relative to the C-3′-OH.

Table 18-1
Ionization Constants of the Ribonucleotides (Presented as p*K* Values)

	Base	Secondary Phosphate	Primary Phosphate
Adenosine-5′-phosphate (5′-AMP)*	3.8	6.1	0.9
Uridine-5′-phosphate (5′-UMP)	9.5	6.4	1.0
Cytidine-5′-phosphate (5′-CMP)	4.5	6.3	0.8
Guanine-5′-phosphate (5′-GMP)	2.4, 9.4	6.1	0.7

*5′-AMP (or 5′-rAMP) refers to the ribonucleotide. The comparable deoxyribonucleotide (deoxynucleotide) is indicated by the symbol 5′-dAMP.

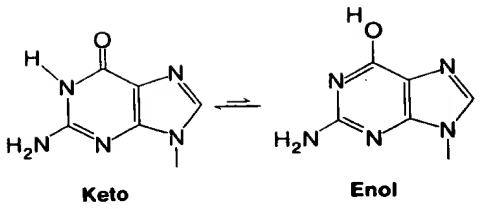


Figure 18-2
Keto-enol isomerization of guanine.

Compositional variability in the nucleotide is accounted for solely by the purine or pyrimidine base attached to the C-1′ position of the sugar (atoms in the sugar are commonly given a prime designation). These bases are either the purines, *adenine* and *guanine*, or the pyrimidines, *thymine* and *cytosine*. The same bases are found in RNA, except that thymine is replaced by *uracil*, which has an H— group instead of a CH_3— group on the C-5 position of the pyrimidine. The sugar in ribonucleotides is also different in having an additional HO group on the C-2′ which is cis with respect to the C-3′-OH.

All the commonly occurring nucleosides and nucleotides are capable of existing in two tautomeric forms. For example, guanosine (G) can undergo the keto-enol shift shown in Figure 18-2. The keto form is strongly favored, so much so that it is difficult to detect even trace amounts of the enol form. Similarly, the keto forms of thymidine (T) or uridine (U) are strongly preferred. Adenosine (A) and cytidine (C) can isomerize to imino forms (not shown), but once again the amino forms (shown) are strongly preferred. Even though the unusual tautomers are present in very small amounts, it is conceivable that they are contributors to the mutation process.

Some nucleotides undergo protonation in acid and some undergo deprotonation in base; the relevant p*K*s are indicated in Table 18-1. At neutrality there is no charge on any of the bases. Three of the bases undergo protonation as the pH is lowered (A, C, and G). X-ray diffraction and spectroscopic analysis (nuclear magnetic resonance and infrared spectroscopy) have been used to show that adenosine protonates on the N-1 position of the purine rather than on the amino group (see Figure 18-3). The charged form is stabilized by the resonance hybrids shown. On cytidylic

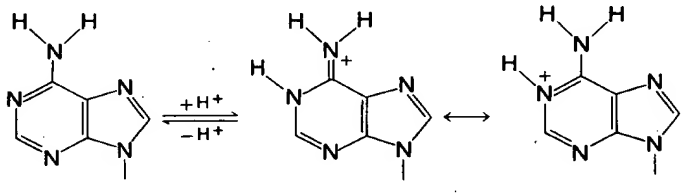


Figure 18-3
Uncharged and protonated forms of adenosine. The charged base resonates between the two structures shown on the right.

polymerases
omotor?

J. D. WATSON
Harvard University and
Cold Spring Harbor Laboratory

MOLECULAR BIOLOGY OF THE GENE

SECOND EDITION

With illustrations by Keith Roberts

W. A. BENJAMIN, INC.
Menlo Park, California • Reading, Massachusetts
London • Amsterdam • Don Mills, Ontario • Sydney

MOLECULAR BIOLOGY OF THE GENE, ***Second Edition***

Standard Book Number 8053-9020-0 (Clothbound Edition)
8053-9603-9 (Paperback Edition)

Library of Congress Catalog Card Number 72-134173

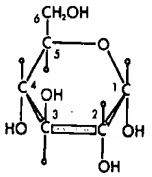
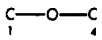
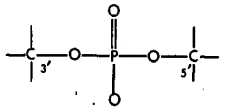
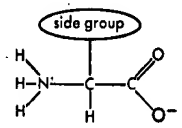
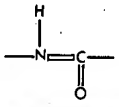
Manufactured in the United States of America

ISBN 0-8053-9602-0 (hardbound)
ISBN 0-8053-9603-0 (paperback)
GHIJKLMNOP-HA-79876543

W. A. BENJAMIN, INC., Menlo Park, California

Frontispiece: *Electron micrograph of polyribosomes attached to a section of the E. coli genome. (Kindly supplied by O. L. Miller, Jr., Biology Division, Oak Ridge National Laboratory, Barbara A. Hamkalo, and C. A. Thomas, Jr., Department of Biological Chemistry, Harvard Medical School.)*

TABLE 3–4 *Structural organization of several important biological macromolecules*

Macromolecule	Monomeric units	Number of different monomers	General monomer formula	Fixed or irregular chain length	Linkage between monomers
Glycogen (a polysaccharide)	Glucose	One		Indefinite—may be > 1000	1–4-Glycosidic linkage 
DNA (deoxyribonucleic acid)	Deoxynucleotides	Four: deoxyadenylate, deoxyguanylate, deoxythymidylate, deoxycytidylate	Purine-deoxyribose-P (or pyrimidine-deoxyribose-P)	Genetically fixed—may be > 10^7	3′–5′-Phosphodiester linkage 
RNA (ribonucleic acid)	Ribonucleotides	Four: adenylate, guanylate, uridylate, cytidylate	Purine-ribose-P (or pyrimidine-ribose-P)	Genetically fixed, often > 3000	3′–5′-Phosphodiester linkage
Protein	L-Amino acids	Twenty: glycine, alanine, serine, etc.		Genetically fixed, usually varies between 100 and 1000	Peptide linkage 

A chemist's lo

of enthusiasm
pure chemistry
when they un
meric molecule
exist well-defin
(metabolic pa
number of enz
in a cell by the

**MACROMOLE
BY LINEAR**

Most of the m
other cells, is
large molecule
enzymes. Th
construct the
that the numb
the number in
that most bi
directly conce
understanding
Furthermore,
size would lea
however, the
zations reduce
task.

First, all ma
the condensat
thesis thus oc
smaller subun
these subunits
house from pr

Second, the
common chen
describes the
macromolecul
tion of nitro

Biochemistry

Reginald H. Garrett

Charles M. Grisham

University of Virginia

SAUNDERS COLLEGE PUBLISHING
HARCOURT BRACE COLLEGE PUBLISHERS

Fort Worth • Philadelphia • San Diego • New York • Orlando • Austin
San Antonio • Toronto • Montreal • London • Sydney • Tokyo

Text Typeface: Baskerville
Compositor: York Graphic Services, Inc.
Acquisitions Editor: John J. Vondeling
Developmental Editor: Sandra Kiselica
Managing Editor: Carol Field
Project Editors: Becca Gruliow, Sarah Fitz-Hugh
Copy Editor: Zanae Rodrigo
Manager of Art and Design: Carol Bleistine
Art Director: Carol Bleistine, Anne Muldrow
Art Assistant: Sue Kinney
Text Designer: Rebecca Lemna
Cover Designer: Lawrence R. Didona
Text Artwork: J/B Woolsey Associates
Layout Artwork: Claudia Durrell
Director of EDP: Tim Frelick
Production Manager: Charlene Squibb
Marketing Manager: Marjorie Waldron
Field Product Manager: Laura Coaty

Cover Credit: J/B Woolsey Associates

Printed in the United States of America

5 6 7 8 9 0 1 2 3 032 10 9 8 7 6 5 4 3

ISBN 0-03-009758-4

Library of Congress Catalog Card Number: 93-087782

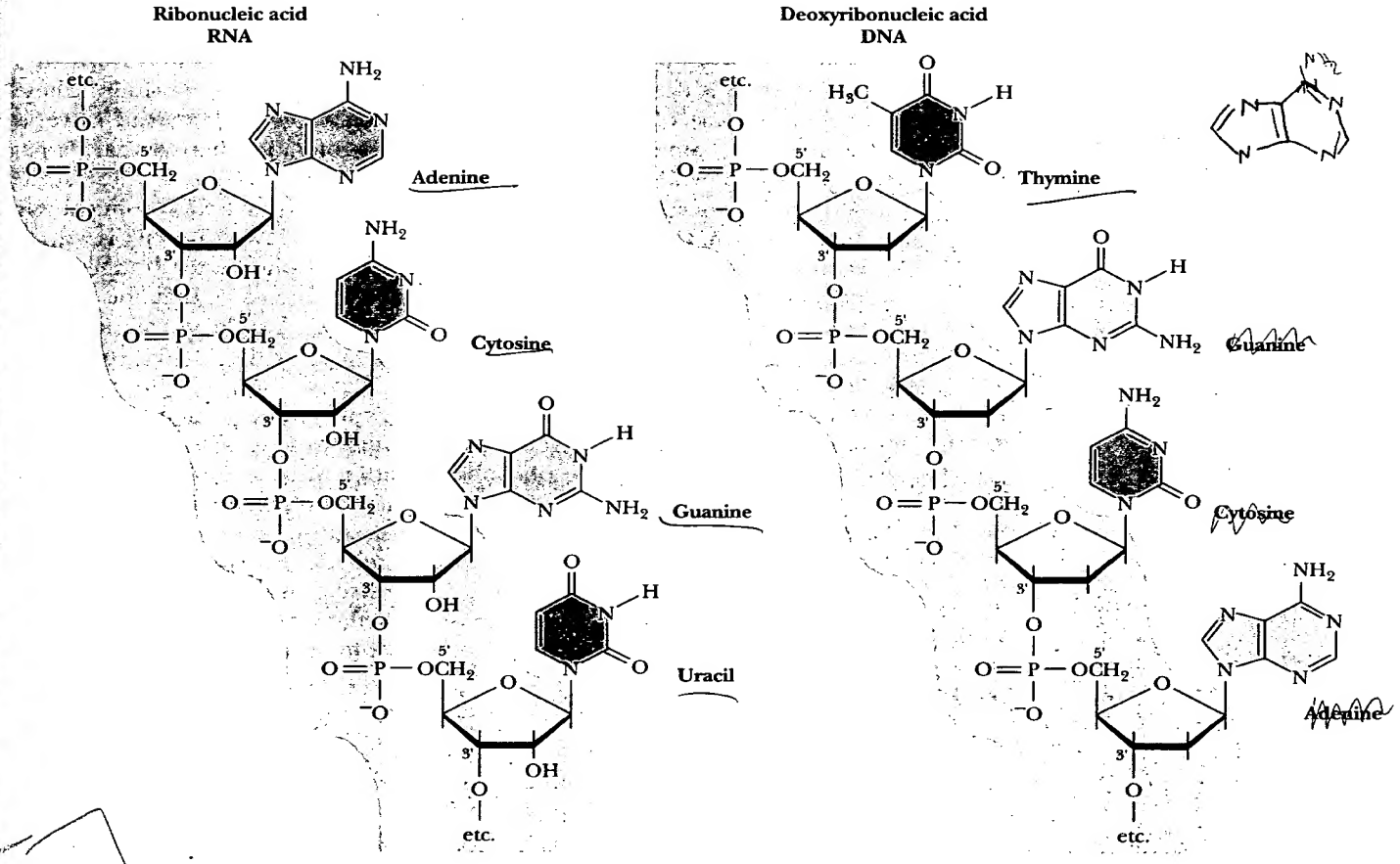


Figure **6.17** 3′-5′ phosphodiester bridges link nucleotides together to form polynucleotide chains.

cleotides, in which the chain can be visualized as running from 5′ to 3′ along the atoms of one furanose and thence across the phosphodiester bridge to the furanose of the next nucleotide in line. Thus, this backbone can be portrayed by the symbol of a vertical line representing the furanose and a slash representing the phosphodiester link, as shown in Figure 6.18. The diagonal slash runs from the middle of a furanose line to the bottom of an adjacent one to indicate the 3′- (middle) to 5′- (bottom) carbons of neighboring furanoses joined by the phosphodiester bridge. The base attached to each furanose is indicated above it by a one-letter designation: A, C, G, or U (or T). The convention in all notations of nucleic acid structure is to read the polynucleotide chain from the 5′-end of the polymer to the 3′-end. Note that this reading direction actually passes through each phosphodiester from 3′ to 5′.

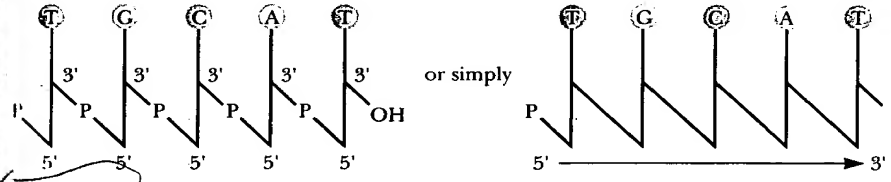
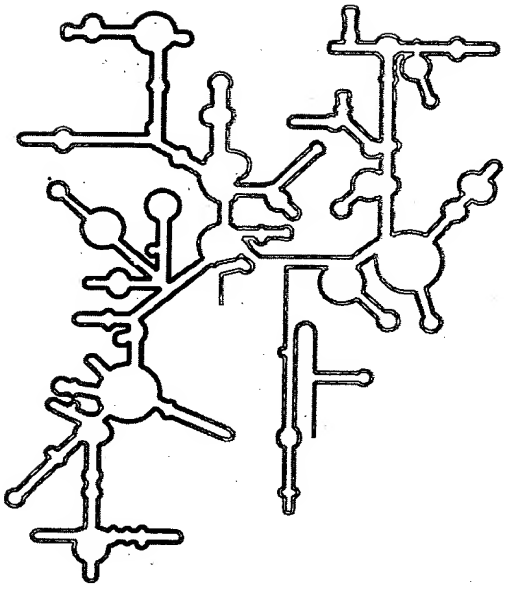


Figure **6.18** Furanoses are represented by lines; phosphodiesters are represented by diagonal slashes in this shorthand notation for nucleic acid structures.

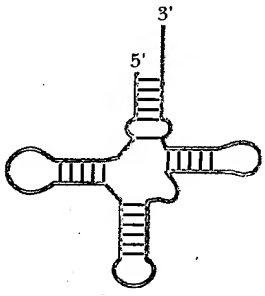
Base Sequence

The only significant variation that commonly occurs in the chemical structure of nucleic acids is the nature of the base at each nucleotide position. These bases are not part of the sugar–phosphate backbone but instead serve as distinctive side chains, like the R groups attached to a polypeptide backbone. They give the polymer its unique identity. A simple notation of these structures is merely to list the order of bases in the polynucleotide using single capital letters—A, G, C, and U (or T). Occasionally, a lowercase "p" is written between each successive base to indicate the phosphodiester bridge, as in GpApCpGpUpA. A "p" preceding the sequence indicates that the nucleic acid carries a PO_4 on its 5′-end, as in pGpApCpGpUpA; a "p" terminating the sequence connotes the presence of a phosphate on the 3′-OH end, as in GpApCpGpUpAp.

A more common method of representing nucleotide sequences is to omit the "p" and write only the order of bases, such as GACGUA. This notation assumes the presence of the phosphodiesters joining adjacent nucleotides. The presence of 3′- or 5′-phosphate termini, however, must still be specified, as in GACGUAp for a 3′-PO_4 terminus. To distinguish between RNA and DNA sequences, DNA sequences are typically preceded by a lowercase "d" to denote deoxy, as in d-GACGTA. From a simple string of letters such as this, any biochemistry student should be able to draw the unique chemical structure for a pentanucleotide, even though it may contain over 200 atoms.



Ribosomal RNA.



tRNA.

6.6 Classes of Nucleic Acids

The two major classes of nucleic acids are DNA and RNA. DNA has only one biological role, but it is the more central one. The information to make all the functional macromolecules of the cell (even DNA itself) is preserved in DNA and accessed through transcription of the information into RNA copies. Coincident with its singular purpose, there is only a single DNA molecule (or "chromosome") in simple life forms such as viruses or bacteria. Such DNA molecules must be quite large in order to embrace enough information for making the macromolecules necessary to maintain a living cell. The *Escherichia coli* chromosome has a molecular mass of 2.9×10^9 daltons and contains 9.5 million nucleotides. Eukaryotic cells have many chromosomes, and DNA is found principally in two copies in the diploid chromosomes of the nucleus, but it also occurs in mitochondria and in chloroplasts, where it encodes a restricted set of proteins and RNAs unique to these organelles.

In contrast, RNA occurs in multiple copies and various forms (Table 6.2). Cells contain up to eight times as much RNA as DNA. RNA has a number of important biological functions, and on this basis, RNA molecules are categorized into several major types: **messenger RNA, ribosomal RNA,** and **transfer RNA.** Eukaryotic cells contain an additional type, **small nuclear RNA (snRNA)** (see Chapter 30). Messenger RNA (**mRNA**) serves to carry the information or "message" that is encoded in genes to the sites of protein synthesis in the cell, where this information is translated into a polypeptide sequence. Because mRNA molecules are transcribed copies of the protein-coding genetic units that comprise most of DNA, it is said to be "the DNA-like RNA." Ribosomes, the supramolecular assemblies where protein synthesis occurs, are 65% RNA of the ribosomal RNA type. Ribosomal RNA (**rRNA**) molecules fold into characteristic secondary structures as a consequence of intramolecular hydrogen bond interactions. There are three major species of rRNA in ribosomes, and they are generally referred to according to their sedimentation coefficients

MODERN CONCEPTS IN
BIOCHEMISTRY

FOURTH EDITION

Robert C. Bohinski
John Carroll University

ALLYN AND BACON, INC.
Boston London Sydney Toronto

The cover illustration is a computer-drawn representation of the Semliki Forest virus. The virus was first isolated in 1944 from the tissues of the mosquito and was named for the rain forest in southern Uganda where the mosquitoes were found.

The virus itself is 650 angstrom units in diameter. The outer viral membrane is composed of two layers of lipid molecules, which are free to move laterally, giving the membrane the properties of a liquid. Inserted into the membrane are 180 "spikes," each made up of three linked protein molecules. Each of the three spike proteins has carbohydrate side chains attached; such proteins are called glycoproteins.

Library of Congress Cataloging in Publication Data

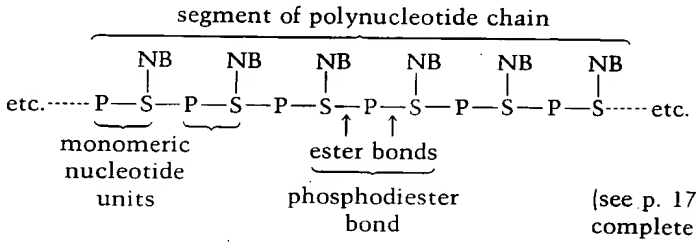
Bohinski, Robert C.
Modern concepts in biochemistry.

Includes bibliographies and index.
1. Biological chemistry. I. Title.
QP514.2.B63 1983 574.19'2 82-16422
ISBN 0-205-07905-9
ISBN (international) 0-205-07968-7

10 9 8 7 6 5 4 3 2 1 87 86 85 84 83

6–1 NUCLEOTIDES—THE UNIT COMPONENTS OF NUCLEIC ACIDS

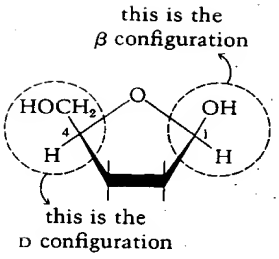
Both RNA and DNA are composed of monomeric units called *nucleotides;* hence a nucleic acid is also a *polynucleotide.* A single nucleotide consists of three chemical parts—inorganic phosphate, a simple sugar, and either a purine or a pyrimidine (called a *nitrogen base*). These are attached to each other in the order: phosphate—sugar—nitrogen base (P,S, and NB in the diagram). Successive nucleotides are linked via an *ester* bond between the sugar and phosphate of adjacent nucleotides. Since the sugar and phosphate within a nucleotide are also linked via an ester bond, the S—P—S linkage along the backbone of a polynucleotide chain is called a *phosphodiester bond.*



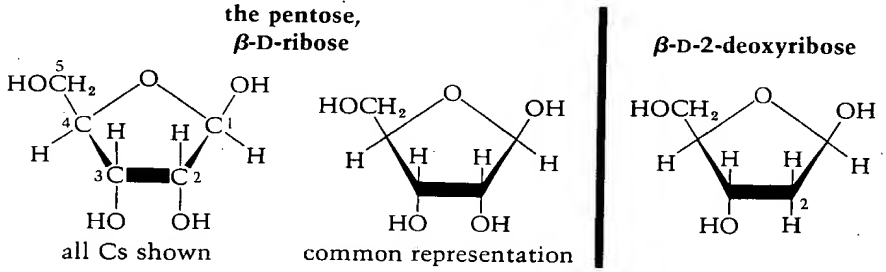
The nitrogen bases are not involved in any covalent linkages other than their attachment to the sugar—phosphate backbone. It is the *sequence of nitrogen bases* along the invariant sugar—phosphate backbone that *constitutes the unique structural and functional individuality of DNA and RNA molecules.* The terms *nucleotide sequence* and *nitrogen base sequence* are used interchangeably.

Let us now examine the specifics of this preliminary description of nucleotide structure. In addition some other important aspects of nucleotide biochemistry will also be covered.

Ribose and Deoxyribose



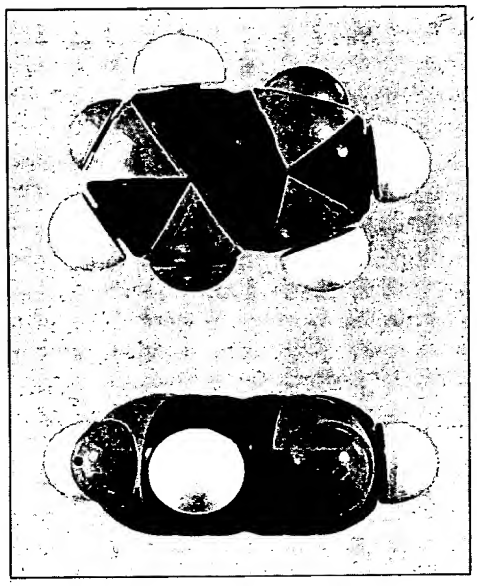
A complete description of sugar structures is given in the next chapter. The objective here is merely to show the structures of the two simple sugars characteristic of RNA or DNA. The nucleotides of RNA contain *β-D-ribose;* the nucleotides of DNA contain *β-D-2-deoxyribose.* Both are pentose sugars (five carbon atoms) differing in structure only at carbon-2. Carbon-2 in ribose has an —OH while carbon-2 in deoxyribose has just —H, hence "deoxy." The *β* and D symbols refer to the particular configurations at C^1 and C^4 in the ring. For now, it is not essential to understand why. These and other points are dealt with in Chap. 7. Shortly we will identify

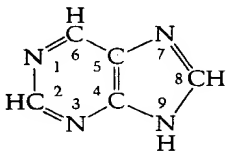


how the nitrogen bases are linked at the C^1 position and how the C^5 and C^3 OH groups participate in phosphoester bonds.

Nitrogen Bases (Purines and Pyrimidines)

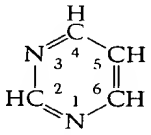
The nitrogen bases found most frequently in RNA and DNA are the *purines: adenine* (A) and *guanine* (G), and the *pyrimidines: cytosine* (C), *thymine* (T), and *uracil* (U). Their structures, systematic names, and occurrence in nucleic acids are given below. The space-filling models clearly illustrate the coplanar bonding pattern in these aromatic rings. Although *keto-enol* forms (*tautomerism*) are possible for G, C, T, and U, the keto structures are much more stable and thus predominate under physiological conditions.



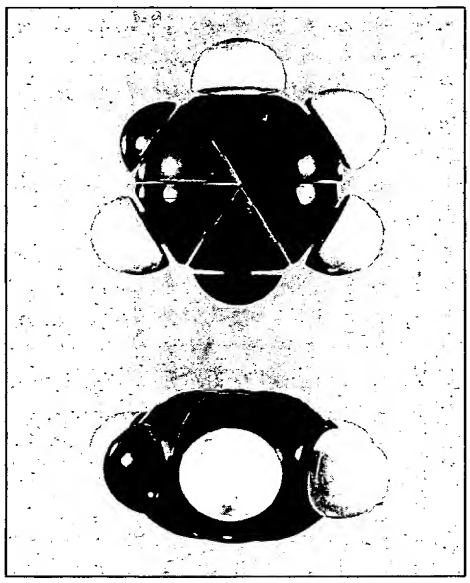


purine

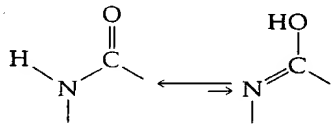
purines and pyrimidines are flat molecules



pyrimidine

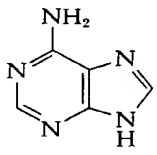


tautomerism:

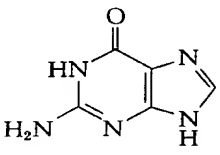


keto form enol form

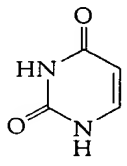
Major purines and pyrimidines of nucleic acids



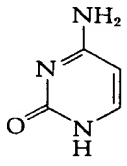
adenine (A)
6-aminopurine
(RNA + DNA)



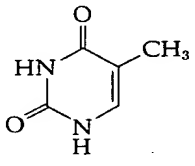
guanine (G)
2-amino-6-oxypurine
(RNA + DNA)



uracil (U)
2,4-dioxypyrimidine
(RNA)



cytosine (C)
2-oxy-4-amino
pyrimidine
(RNA + DNA)



thymine (T)
5-methyl-2,4-
dioxypyrimidine
(DNA; also
found in
transfer-RNA)